AD-A080 264 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMA--ETC F/G 12/1
SOME ALGORITHMS FOR THE RECURSIVE INPUT-OUTPUT MODELING OF 2-D --ETC(U)
DEC 79 B C LEVY, M MORF, S KUNG MDA903-78-C-0179
UNCLASSIFIED LIDS-P-962 AFOSR-TR-80-0070 NL

1 OF 1
AD A 080264

END
DATE FILMED
3-80
DDC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ADA080264

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR TR-80-0070
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): SOME ALGORITHMS FOR THE RECURSIVE INPUT-OUTPUT MODELING OF 2-D SYSTEMS
5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Bernard C. Levy, Sun-Yuan Kung, Martin Morf
8. CONTRACT OR GRANT NUMBER(s): MDA903-78-C-0179, AFOSR-77-3281
9. PERFORMING ORGANIZATION NAME AND ADDRESS: M.I.T. Laboratory for Information & Decision Systems, Cambridge, MA 02139
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, D.C. 20332
12. REPORT DATE: December 1979
13. NUMBER OF PAGES: 56
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

LEVEL II

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

LIDS-P-962

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

2304

A1

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

410 150

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers the deterministic and stochastic modeling of 2-D systems described by their input/output data. In the deterministic case, the modeling problem is formulated as a 2-D Pade approximation problem. By studying several possible geometries of approximation, we obtain several sets of recursions of the 2-D rational approximants. These resutls exploit the properties of 2-D Hankel matrices, and they are used here to characterize the 2-D rational transfer functions. In the stochastic case, the realization problem is viewed as a 2-D prediction problem. This problem is solved recursively by generalizing to the

DD FORM 1473

Unclassified

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract cont.

2-D case an algorithm due to Levinson in the 1-D case. The predictors obtained by this algorithm are then showed to converge to the 2-D spectral factors of the output spectrum.

UNCLASSIFIED

AFOSR-TR- 80-0070

December 1979 LIDS-P-962

SOME ALGORITHMS FOR THE RECURSIVE INPUT-OUTPUT MODELING OF 2-D SYSTEMS†

Bernard C. Lévy*
Martin Morf**
Sun-Yuan Kung***

ABSTRACT

This paper considers the deterministic and stochastic modeling of 2-D systems described by their input/output data. In the deterministic case, the modeling problem is formulated as a 2-D Padé approximation problem. By studying several possible geometries of approximation, we obtain several sets of recursions of the 2-D rational approximants. These results exploit the properties of 2-D Hankel matrices, and they are used here to characterize 2-D rational transfer functions. In the stochastic case, the realization problem is viewed as a 2-D prediction problem. This problem is solved recursively by generalizing to the 2-D case an algorithm due to Levinson in the 1-D case. The predictors obtained by this algorithm are then showed to converge to the 2-D spectral factors of the output spectrum.

†This work was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-77-3281 and in part by the Defense Advanced Project Agency under contract MDA903-78-C-0179.

*Laboratory for Information and Decision Systems and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

**Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

***Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007.

I. Introduction

In this paper, the modeling of 2-D linear and shift-invariant systems given by their input/output data will be considered. This problem will be studied both in the deterministic case, i.e. when the point spread function of the system is known, and in the stochastic case when only input and output covariances are given. For convenience, we shall restrict our attention to the case of single input-single output systems.

In the deterministic case, the realization problem will be formulated as a 2-D Padé approximation problem. Thus, given a point spread function

$$g(z,\omega) = \sum_{i,j} g_{ij}\, z^{-i}\omega^{-j} \tag{1}$$

we shall approximate g by rational functions

$$\hat{g}(z,\omega) = \frac{b(z,\omega)}{a(z,\omega)} \tag{2}$$

where a and b are 2-D polynomials. Such a problem has been considered by Chisholm [1 Hughes Jones and Makinson [2], Graves-Morris, Hughes Jones and Makinson [3] among oth but the results discussed here will be somewhat different. The main difference is that the realization algorithms presented in Sections II and III are recursive. These algorithms extend to the 2-D case a realization procedure originally introduced by Lanczos [4], Berlekamp [5], Massey [6] and Rissanen [7] in the 1-D case. The recursions that we obtain exploit the properties of Hankel block Hankel matrices. These properties can be used to relate the recursions obtained for $\hat{g}$ to those derived by Jackson [8] for 2-D orthogonal polynomials on the hyper real line.

In Section IV we study the relation existing between the 2-D partial realization problem considered here and the complete realization problem studied by Ho and Kalman [9], Silverman [10] and Youla and Tissi [11] in the 1-D case, and by

Fliess [12], Fornasini [13], Clerget [14] and Kao and Chen [15] in the 2-D case. By doing so, we obtain a set of simple conditions on the Markov parameters of g that can be used to verify whether the partial realization $\hat{g} = b/a$ is also a complete realization, i.e. whether $\hat{g} = g$.

The stochastic realization problem is considered in Section V and is formulated as an autoregressive modeling problem. Thus, if the system considered is linear and shift invariant, and if it is driven by a 2-D white noise process u(i,j), the output process will be modeled as

$$y(i,j) + \sum_{I-(0,0)} a_I(k,\ell)\, y(i-k,j-\ell) = u(i,j) \tag{3}$$

where I is a causal asymetric half plane set of the type considered in [16]-[18]. The coefficients $a_I(k,\ell)$ of the filter (3) will be selected so that

$$\hat{y}(i,j) = -\sum_{I-(0,0)} a_I(k,\ell)\, y(i-k,j-\ell) \tag{4}$$

is the linear least-squares estimate of y(i,j) given observations over the set

$$I(i,j) = \{(i-k,j-\ell) : (k,\ell) \in I - (0,0)\}.$$

To solve this linear least-squares estimation problem, we shall use a 2-D version of an algorithm originally introduced by Levinson [19] in the 1-D case. The recursions that we obtain for $a_I(.,.)$ were first described in [20], [21], and they differ from those presented by Justice [22] by the fact that the orders n and m of a_I in z and ω can be increased separately (Justice's recursions were requiring that either n or m be fixed a priori). These recursions present also some similarity with the recursions derived by Genin and Kamp [23] for 2-D orthogonal polynomials on the unit hypercircle (see [17]-[18] and [24]-[27] for more details).

In addition, we relate the stochastic modeling problem to the spectral factorization results of Helson and Lowdenslager [28], Pistor [29], Ekstrom and Woods [16], Murray [30] and Genin and Kamp [18]. It is shown in this context that when the domain I becomes infinite, $a_I(z,\omega)$ converges to the spectral factor of the output spectrum $r(z,\omega)$. Finally, the Section VI contains some observations on some open problems and on possible extensions of the results discussed here.

II. The 2-D Padé Approximation Problem

Throughout the following sections, it will be assumed that the 2-D transfer function $g(z,\omega)$ that we want to realize is a South-West (SW) quarter plane causal filter, i.e.

$$g(z,\omega) = \sum_{i\geq 0,\ j\geq 0} g_{ij}\ z^{-i}\omega^{-j} \tag{5}$$

so that $g(z,\omega)$ will be modeled by a transfer function $\hat{g} = b/a$ such that

$$a(z,\omega) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_{ij}\ z^{n-i}\omega^{m-j}\ , \quad a_{00} = 1 \tag{6a}$$

is monic, and

$$b(z,\omega) = \sum_{i=0}^{n}\sum_{j=0}^{m} b_{ij}\ z^{n-i}\omega^{m-j}\ . \tag{6b}$$

There is no loss of generality in making the previous assumption since an arbitrary transfer function $g(z,\omega)$ can always be decomposed into four parts which are causal in each of the four quadrants, i.e.

$$g = g_{SW} + g_{NW} + g_{SE} + g_{NE}.$$

These four parts can then be approximated separately.

In the following, we shall consider two types of Padé approximants for g.

Definition Let $\hat{g}(z,\omega) = b(z,\omega)/a(z,\omega)$ be a 2-D rational function. Then $\hat{g}$ is said to be a rational Padé approximant of g in the domain $D \subset \mathbb{N}^2$ if

$$g(z,\omega) - \hat{g}(z,\omega) = \sum_{(i,j)\,\varepsilon\,\bar{D}} d_{ij} z^{-i} \omega^{-j} \qquad (7)$$

where $\bar{D}$ is the complement of D. Similarly, $\hat{g}$ is said to be a modified Padé approximant of g over the domain $\Delta \subset \mathbb{Z}^2$ if

$$a(z,\omega)\; g(z,\omega) - b(z,\omega) = \sum_{(i,j)\,\varepsilon\,\bar{\Delta}} \delta_{ij} z^{-i} \omega^{-j} \qquad (8)$$

where $\bar{\Delta}$ is the complement of Δ.

The motivation for making the distinction between the problems (7) and (8) is that it is not always possible to convert a rational approximation problem into a modified one and vice-versa. An example of domain of modified approximation that cannot be converted into an equivalent domain of rational approximation is described in Figure 1. In this case, the parameters of $a(z,\omega)$ and $b(z,\omega)$ satisfy $2(n+1)(m+1)$ linear constraints, but if we consider $\hat{g} = b/a$, the Markov parameters g_{ij} of $g(z,\omega)$ are matched by those of $\hat{g}$ only over the domain

$$D(n,m) = \{(i,j) : (0,0) \le (i,j) \le (n,m)\}$$

which contains $(n+1)(m+1)$ parameters. The modified approximation problem associated to Figure 1 will be considered in Section III. However, we shall consider also several geometries of rational approximation that can be converted easily into modified ones. The Figure 2 describes several cases that will be discussed below.

To justify the study of both the problems (7) and (8), it is useful to note that when $\hat{g} = b/a$ is rational, the rational approximant $\hat{g}_R$ and the modified approximant $\hat{g}_M$ are such that

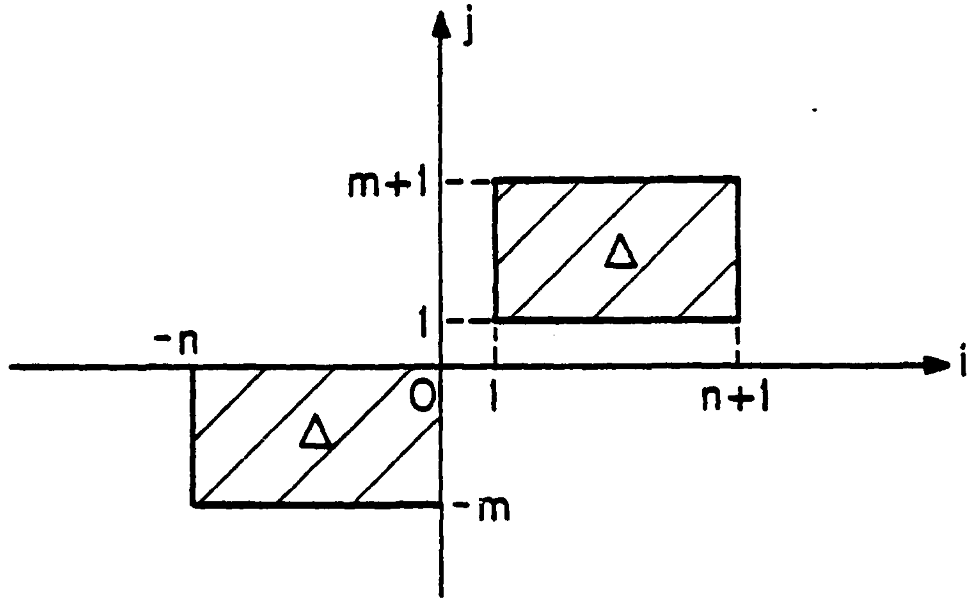

Figure 1 A Geometry of Modified Approximation

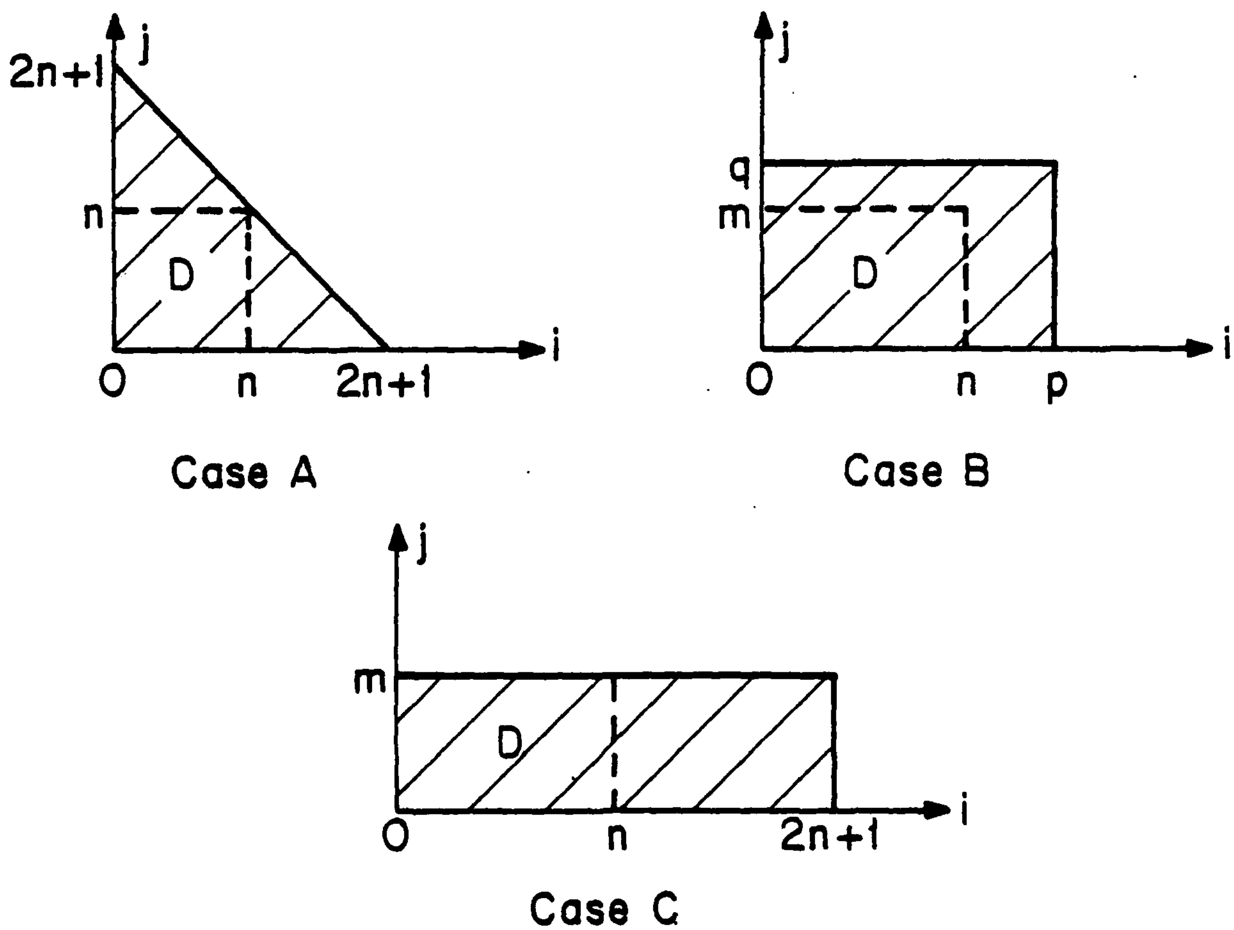

Figure 2 Some Geometries of Rational Approximation

$$\hat{g}_R = \hat{g}_M = b/a$$

provided that the domains D and Δ are chosen sufficiently large. In the remainder of this section we shall discuss the rational approximation problem for the cases A-C of Figure 2.

Case A: This case was the one considered by Chisholm [1], Hughes Jones and Makinson [2] and Graves-Morris, Hughes Jones and Makinson [3]. In this case, the Padé approximants have several interesting properties (such as invariance under a large class of transformations), but this geometry does not permit the efficient exploitation of the shift-invariance properties of the realization problem. This means that the Padé approximants $\hat{g}_{n,n}(z,\omega)$ where

$$n = \deg_z a = \deg_\omega a$$

cannot be computed recursively.

Case B: If the domain of rational approximation is given by

$$D(p,q) = \{(i,j) : (0,0) \leq (i,j) \leq (p,q)\} \tag{9}$$

and if we assume that

$$(n = \deg_z a,\ m = \deg_\omega a) \leq (p,q)$$

then the rational approximation problem (7) can be transformed into an equivalent modified approximation problem where the domain Δ of approximation is given by

$$\Delta(n,m;\ p,q) = \{(i,j) : (-n,-m) \leq (i,j) \leq (p-n,\ q-m)\}. \tag{10}$$

This means that if we denote by Π_Δ the projection operator which selects the co-

efficients of a Laurent power series

$$h(z,\omega) = \sum_{Z^2} h_{ij}\, z^{-i}\omega^{-j}$$

which belong to Δ, i.e.

$$\Pi_\Delta h(z,\omega) = \sum_{\Delta} h_{ij}\, z^{-i}\omega^{-j} ,$$

then $a(z,\omega)$ and $b(z,\omega)$ satisfy the equation

$$\Pi_{\Delta(n,m;\ p,q)}\, a(z,\omega)\ g(z,\omega) - b(z,\omega) = 0 . \tag{11}$$

Consequently, the coefficients of $a(z,\omega)$ and $b(z,\omega)$ obey a set of linear equations given by

$$b_{ij} = \sum_{D(n,m)} g_{i-k,j-\ell}\, a_{k\ell},\ (i,j) \in D(n,m) \tag{12}$$

$$0 = \sum_{D(n,m)} g_{i-k,j-\ell}\, a_{k\ell},\ (i,j) \in D(p,q) - D(n,m) . \tag{13}$$

By scanning the set $D(p,q)$ <u>row by row</u> and by denoting

$$a_j(n;p) = \begin{pmatrix} a_{oj} \\ a_{ij} \\ \underline{a_{nj}} \\ 0_{p-n} \end{pmatrix} \qquad b_j(n;p) = \begin{pmatrix} b_{oj} \\ b_{ij} \\ \underline{b_{ng}} \\ 0_{p-n} \end{pmatrix} \qquad 0 \le j \le m \tag{14a}$$

$$G_j(p) = \begin{pmatrix} g_{oj} & & & 0 \\ g_{ij} & \ddots & & \\ & \ddots & \ddots & \\ g_{pj} & & g_{ij} & g_{oj} \end{pmatrix} \qquad (14b)$$

where 0_{p-n} is a zero $p-n \times 1$ vector, we can rewrite (12) and (13) in matrix form. Indeed, if $\otimes$ denotes the matrix Kronecker product, and if we define

$$a(n,m;\ p,q) = \begin{pmatrix} a_o(n;p) \\ \vdots \\ a_m(n;p) \\ \hline 0_{q-m} \otimes 0_{p+1} \end{pmatrix} \quad b(n,m;\ p,q) = \begin{pmatrix} b_o(n;p) \\ \\ b_m(n;p) \\ \hline 0_{q-m} \otimes 0_{p+1} \end{pmatrix}$$

$$G_{(p,q)} = \begin{pmatrix} G_o(p) & & & 0 \\ G_1(p) & \ddots & & \\ & \ddots & \ddots & \\ G_q(p) & & G_1(p) & G_o(p) \end{pmatrix} \qquad (15)$$

it is easy to see that (12) and (13) are equivalent to

$$b(n,m;\ p,q) = G(p,q)\ a(n,m;\ p,q)\ . \qquad (16)$$

This equation has several interesting features. One of them is that $G(p,q)$ is a block lower triangular Toeplitz matrix whose blocks are themselves lower triangular Toeplitz. This property will be denoted as $G(p,q) \in LT^2$, and it will be shown below that this structure has several advantages. Another interesting aspect

of (16) is that the coefficients of $b(z,\omega)$ are entirely given by those of $a(.,.)$, so that the main problem is to compute $a(.,.)$.

To solve (16), one needs to match the $(p+1)(q+1)$ elements of $\Delta(n,m;\ p,q)$ with the $2(n+1)(n+1)$ coefficients of $a(.,.)$ and $b(.,.)$, so that we have to select p and q such that $2(n+1)(m+1) \geq (p+1)(q+1)$. One such choice is given by

$$p = 2n+1 \quad , \quad q = m \tag{17a}$$

or symmetrically by

$$p = n \quad , \quad q = 2m+1 \tag{17b}$$

This choice corresponds to the geometry C of Figure 2. Other choices are possible, but they do not seem to give rise to recursions in n and m for the polynomials $a(.,.)$ and $b(.,.)$. Another approach that will be considered in Section III is when

$$p = 2n+1 \quad , \quad q = 2m+1 \tag{18}$$

and when only the parameters of the subset of $\Delta(n,m;\ 2n+1,\ 2m+1)$ described in Figure 1 are matched. As we shall see, this case is the most closely related to the 1-D case (it depends on the properties of 2-D Hankel matrices).

<u>Case C</u>: If p and q are selected as in (17a), we can denote by

$$\hat{g}_{n,m}(z,\omega) = \frac{b_{n,m}(z,\omega)}{a_{n,m}(z,\omega)}$$

the Padé approximant that matches the Markov parameters g_{ij} in the domain $D(2n+1,m)$. Since $a_{n,m}(z,\omega)$ is chosen to be <u>monic</u> (i.e. its leading coefficient $a_{n,m}(0,0)=1$), the number of free coefficients of $\hat{g}_{n,m}(z,\omega)$ is only $2(n+1)(m+1)-1$, which is one

less than the number of parameters in $D(2n+1, m)$ or in $\Delta(n,m;\ 2n+1, m)$. Therefore, it will be assumed in the following that the coefficient of $z^{-(n+1)}\omega^m$ in $\Delta(n,m);\ 2n+1, m)$ is <u>not matched</u>. In this case, (11) becomes

$$\Pi_{\Delta(n,m;\ 2n+1,\ m)}\ a_{n,m}\ (z,\omega)\ g(z,\omega) = b_{n,m}\ (z,\omega) + \delta_{n,m}\ z^{-(n+1)}\omega^m \tag{19}$$

where $\delta_{n,m}$ is the residual corresponding to the mis-match of $z^{-(n+1)}\omega^m$.

Since $\Delta(n,m;\ 2n+1,\ m) = -D(n,m) \cup \Sigma(n,m)$, where

$$-D(n,m) = \{(i,j)\ :\ (-i,-j)\ \varepsilon\ D(n,m)\}$$

$$\Sigma(n,m) = \{(i,j)\ :\ (i-1,-j)\ \varepsilon\ D(n,m)\}$$

we can decompose (19) into two parts:

$$\Pi_{-D(n,m)}\ a_{n,m}\ (z,\omega)\ g(z,\omega) = b_{n,m}\ (z,\omega) \tag{20}$$

expresses $b_{n,m}\ (z,\omega)$ in function of $a_{n,m}\ (z,\omega)$, and

$$\Pi_{\Sigma(n,m)}\ a_{n,m}\ (z,\omega)\ g(z,\omega) = \delta_{n,m}\ z^{-(n+1)}\omega^m \tag{21}$$

is the equation satisfied by $a_{n,m}\ (z,\omega)$. To obtain the matrix form of these equations, one can denote by

$$a(n,m) = \begin{pmatrix} a_o(n,m) \\ a_j(n,m) \\ a_m(n,m) \end{pmatrix}, \qquad b(n,m) = \begin{pmatrix} b_o(n,m) \\ b_j(n,m) \\ b_m(n,m) \end{pmatrix} \tag{22}$$

with

$$a_j(n,m) = \begin{pmatrix} a_{n,m}(o,j) \\ a_{n,m}(i,j) \\ a_{n,m}(n,j) \end{pmatrix}, \qquad b_j(n,m) = \begin{pmatrix} b_{n,m}(o,j) \\ b_{n,m}(i,j) \\ b_{n,m}(n,j) \end{pmatrix}.$$

the vectors which are obtained by scanning the coefficients of $a_{n,m}(z,\omega)$ and $b_{n,m}(z,\omega)$ row by row. In this case, (20) can be written

$$b(n,m) = G(n,m)\, a(n,m) \tag{23}$$

where G(n,m) is defined as in (15). Similarly, one can define

$$\tau_j(n) = \begin{pmatrix} g_{n+1\,j} & & g_{ij} \\ & \ddots & \\ g_{2n+1\,j} & & g_{n+1\,j} \end{pmatrix} \tag{24a}$$

and

$$\tau(n,m) = \begin{pmatrix} \tau_0(n) & & & 0 \\ \tau_1(n) & \ddots & & \\ & \ddots & \ddots & \\ \tau_m(n) & & \tau_1(n) & \tau_0(n) \end{pmatrix} \tag{24b}$$

where the matrix $\tau(n,m)$ has a lower triangular Toeplitz block Toeplitz structure. Then, if we introduce

$$\delta(n,m) = \left.\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right\} m+1 \otimes \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_{n,m} \end{pmatrix}\right\} n+1\,, \tag{25}$$

the matrix form of (21) is given by

$$\tau(n,m)\ a(n,m) = \delta(n,m)\ . \tag{26}$$

To compute a(n,m) and b(n,m) recursively, we will now exploit the structure of the matrices G(n,m) and $\tau(n,m)$. This gives the following recursions.

<u>Increase in m</u>

The lower triangular structure of $\tau(n,m)$ can be exploited by assuming that in (22) the vectors $a_j(n,m)$ do not depend on m, i.e.

$$a_j(n,m) \;=\; a_j(n) \quad \text{for all } j. \tag{27}$$

This implies that the vectors $b_j(n,m)$ do not depend on m either. Thus, given a(n,m) and b(n,m), to compute a(n,m+1) and b(n,m+1), we need only to find $a_{m+1}(n)$ and $b_{m+1}(n)$. This can be done by direct substitution, so that

$$a_{m+1}(n) = \tau_0^{-1}(n) \sum_{j=1}^{m} \tau_j(n)\ a_{m+1-j}(n) \tag{28}$$

$$b_{m+1}(n) = \sum_{j+0}^{m+1} G_j(n)\ a_{m+1-j}(n). \tag{29}$$

Since $\tau_0(n)$ is a Toeplitz matrix, its inverse can be computed with $0(n^2)$ operations by using the inversion algorithm of Levinson [19] and Trench [31], or with $0(n\ \log^2 n)$ operations if we use a more efficient version of this algorithm based on doubling ideas (cf. Gustavson and Yun [32], Morf [33] and Bitmead and Anderson [34]). In addition, since the matrices $\tau_j(n)$ and $G_j(n)$ have a Toeplitz structure, the vectors $\tau_j(n)\ a_{m+1-j}(n)$ and $G_j(n)\ a_{m+1-j}(n)$ can be computed by using fast convolution algorithms. This shows that the number of operations required by the recursions (28) and (29) is $0(mn \log n)$.

Increase in n

To compute a(n+1, m), we can use the Toeplitz structure of τ. To do so, we consider the identity

$$\tau(n,m)\;\begin{pmatrix} a_o(n) & & \\ & a_o(n) & \\ a_{m-1}(n) & & \ddots \\ a_m(n) & a_{m-1}(n) & a_o(n) \end{pmatrix} = \begin{pmatrix} \delta(n) & & \\ & \delta(n) & 0 \\ 0 & & \ddots \\ & & \delta(n) \end{pmatrix}$$

$$\underbrace{\quad}_{a(n,m)} \qquad \underbrace{\quad}_{a(n,m-1)} \qquad \underbrace{\quad}_{a(n,0)} \qquad\qquad (30)$$

where

$$\delta(n) \triangleq \left.\begin{pmatrix} 0 \\ | \\ | \\ | \\ 0 \\ \delta_n \end{pmatrix}\right\} n+1 \;^{*}.$$

Then, in order to replace the row by row scanning of the set of D(n,m) by a column by column scanning, we can define the transformation

$$P(n,m) = I_{m+1} \bar{\otimes} I_{n+1}$$

where A $\bar{\otimes}$ B denotes the Paley product of two matrices. If A and B are some matrices of size n $\times$ n and m $\times$ m, this product is defined as

---

* A consequence of the structure (27) for a(n,m) is that the residual $\delta_{n,m}$ does not depend on m.

$$A \bar{\otimes} B = \begin{pmatrix} A \otimes B_1 \\ A \otimes B_j \\ A \otimes B_m \end{pmatrix}$$

where B_j is the j^{th} row of B. By multiplying the rows and columns of $\tau(n,m)$ by $P(n,m)$, and by observing that $P(n,m) = P^T(n,m) = P^{-1}(n,m)$, one can rewrite (30) as

$$\hat{\tau}(n,m)\ \hat{A}(n,m) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hat{\Delta}(n,m) \end{pmatrix} . \tag{31}$$

Here, we have

$$\hat{\tau}(n,m) = \begin{pmatrix} \hat{\tau}_{n+1}(m) & \ddots & \hat{\tau}_1(m) \\ & & \\ \hat{\tau}_{2n+1}(m) & & \hat{\tau}_{n+1}(m) \end{pmatrix} \tag{32a}$$

with

$$\hat{\tau}_i(m) = \begin{pmatrix} g_{i0} & & \\ g_{i1} & g_{i0} & 0 \\ \vdots & \ddots & \ddots \\ g_{im} & g_{i1} & g_{i0} \end{pmatrix} \tag{32b}$$

and

$$\hat{A}(n,m) = \begin{pmatrix} \hat{A}_o(n,m) \\ \hat{A}_i(n,m) \\ \hat{A}_n(n,m) \end{pmatrix} \tag{33a}$$

with

$$\hat{A}_i(n,m) = \begin{pmatrix} a_n(i,0) & & & \\ a_n(i,1) & \ddots & a_n(i,0) & 0 \\ & & \ddots & \\ a_n(i,m) & & a_n(i,1) & a_n(i,0) \end{pmatrix} \tag{33b}$$

where $a_n(i,j) = a_{n,m}(i,j)$ is the $(i,j)^{th}$ coefficient of $a_{n,m}(z,\omega)$, i.e.

$$a_{n,m}(z,\omega) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{n,m}(i,j)\, z^{n-i} \omega^{m-j}$$

(observe that the structure (27) for a(n,m) implies that $a_{n,m}(i,j)$ does not depend on m). Also, the $m + 1 \times m + 1$ matrix $\hat{\Delta}(n,m)$ is given by

$$\hat{\Delta}(n,m) = \mathrm{diag}\ \{\delta_n\}\ . \tag{34}$$

The main aspect of (31) is that $\hat{\tau}(n,m)$ and $\hat{A}(n,m)$ have both lower triangular Toeplitz (LT) block entries. Since the algebra of LT matrices is closed (the product of two LT matrices is commutative and LT, the inverse of a nonsingular LT matrix is LT), one can operate on the blocks $\hat{\tau}_i$ and $\hat{A}_i$ as if they were scalars. This means that to solve (31) efficiently, we will need only to exploit the block Toeplitz structure of $\hat{\tau}(n,m)$. This will be done by using a set of recursions derived originally by Lanczos [4], and introduced in the context of realization theory

by Kung [35] and Kailath [36].

These recursions are based on the observation that we have

$$\hat{\tau}(n+1,m)\left(\begin{array}{c|c|c} & 0 & 0 \\ & \overline{} & 0 \\ \hat{A}(n,m) & & \overline{} \\ & \hat{A}(n,m) & \hat{A}(n-1,m) \\ \overline{} & & \\ 0 & & \end{array}\right) = \left(\begin{array}{c|c|c} & 0 & \\ \hline \hat{\Delta}(n,m) & 0 & \hat{\Delta}(n-1,m) \\ \hat{R}(n,m) & \hat{\Delta}(n,m) & \hat{R}(n-1,m) \\ \hat{S}(n,m) & \hat{R}(n,m) & \hat{S}(n-1,m) \end{array}\right) \tag{35}$$

where $\hat{R}(n,m)$ and $\hat{S}(n,m)$ are LT matrices since they are obtained as the combination of LT matrices. Consequently, if we define

$$\hat{M}(n,m) = \hat{\Delta}^{-1}(n-1,\ m)\,\hat{\Delta}(n,m) \tag{36a}$$

$$\hat{N}(n,m) = \hat{\Delta}^{-1}(n,m)\,\hat{R}(n,m) - \hat{\Delta}^{-1}(n-1,\ m)\,\hat{R}(n-1,\ m)\ , \tag{36b}$$

the matrix

$$\hat{A}(n+1,\ m) = \begin{pmatrix} \hat{A}(n,m) \\ \overline{} \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ \overline{} \\ \hat{A}(n,m) \end{pmatrix} \hat{N}(n,m) - \begin{pmatrix} 0 \\ 0 \\ \overline{} \\ \hat{A}(n-1,\ m) \end{pmatrix} \hat{M}(n,m) \tag{37}$$

will satisfy the equation (31), provided that we replace n by n + 1. In this case, the new residual is given by

$$\hat{\Delta}(n+1,\ m) = \hat{S}(n,m) - \hat{R}(n,m)\,\hat{N}(n,m) - \hat{S}(n-1,\ m)\,\hat{M}(n,m)\ . \tag{38}$$

Since LT matrices form a closed algebra, the recursions (37) involve only the multiplication of $m + 1 \times m + 1$ LT matrices. Therefore, the recursions (37) require only $O(nm \log m)$ operations if one used fast Fourier transform techniques to multiply LT matrices.

Remark The LT matrix $\hat{\Delta}(n+1, m)$ obtained in (38) is not diagonal in general. If one wants $\hat{\Delta}(n+1, m)$ to be diagonal as in (34), one needs only to factor $\hat{\Delta}(n+1, m)$ in its lower triangular part with unit diagonal, times its diagonal part. Then, we can renormalize $\hat{A}(n+1, m)$ accordingly.

Another useful observation is that the residual matrices $\hat{\Delta}(n,m) = \operatorname{diag}\{\delta_n\}$ have to be assumed nonsingular for the previous algorithm to be valid. When $\delta_n = 0$ for some n, a generalized set of recursions (the Berlekamp-Massey recursions) have to be used. These recursions were introduced in the 1-D scalar case by Berlekamp [5] and Massey [6], and then generalized to the 1-D matrix case by Dickinson, Morf and Kailath [37]. However, we will not consider this case here, and it will be assumed that $\delta_n \neq 0$ for all n.

III. A Modified Approximation Problem

The geometry $\Delta(n,m;\ 2n+1,\ m)$ is not the only one for which the modified Padé approximation problem (8) has a recursive solution. In this section, we shall consider the geometry $\Delta(n,m)$ described in Figure 1. As mentioned earlier, the modified approximation problem associated with $\Delta(n,m)$ does not admit an equivalent formulation in terms of rational approximants. However, if $\hat{g}_{n,m}(z,\omega) = b_{n,m}(z,\omega)/a_{n,m}(z,\omega)$ is the modified Padé approximant associated with $\Delta(n,m)$, we shall show that the polynomials $a_{n,m}(z,\omega)$ and $b_{n,m}(z,\omega)$ obey a set of recursions similar to those obtained by Jackson [8] for 2-D orthogonal polynomials on the hyper real line.

The first step is to decompose $\Delta(n,m)$ as $\Delta(n,m) = -D(n,m) \cup D^{+}(n,m)$ with

$$D^{+}(n,m) = \{(i,j) : (i-1,\ j-1) \in D(n,m)\}.$$

Then, by considering the equation

$$\Pi_{\Delta(n,m)}\ a_{n,m}(z,\omega)\ g(z,\omega) - b_{n,m}(z,\omega) = 0 \tag{39}$$

we can verify that $b_{n,m}(z,\omega)$ obeys the same equation as in Case C of Section II (it is given by (20) or (23)). However, $a_{n,m}(z,\omega)$ satisfies a different relation given by

$$\Pi_{D^{+}(n,m)}\ a_{n,m}(z,\omega)\ g(z,\omega) = \delta_{n,m}\ z^{-(n+1)}\ \omega^{-(m+1)} \tag{40}$$

provided that we assume that the coefficient of $z^{-(n+1)}\ \omega^{-(m+1)}$ is not matched. Note that, as in Case C of Section II, the polynomial $a_{n,m}(z,\omega)$ has only $(n+1)(m+1) - 1$ free coefficients, or one less than the number of parameters in $D^{+}(n,m)$. If $a(n,m)$ denotes the vector obtained by scanning the coefficients of $a_{n,m}(z,\omega)$ row by row, (40) can be tranformed into

$$T(n,m)\ a(n,m) = e(n,m) \tag{41}$$

where

$$e(n,m) = \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}\right\} m+1 \otimes \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_{n,m} \end{pmatrix}\right\} n+1 \tag{42}$$

and

$$T(n,m) = \begin{pmatrix} T_{m+1}(n) & & T_1(n) \\ & \ddots & \\ T_{2m+1}(n) & & T_{m+1}(n) \end{pmatrix} \tag{43a}$$

with

$$T_j(n) = \begin{pmatrix} g_{n+1j} & & g_{1j} \\ & \ddots & \\ g_{2n+1j} & & g_{n+1j} \end{pmatrix} . \tag{43b}$$

The matrix T(n,m) is a Toeplitz block Toeplitz matrix, and it corresponds to a simple reordering of the 2-D Hankel matrix H(n,m) associated with $g(z,\omega)$. Indeed, if one denotes by J_k the $k \times k$ matrix given by

$$J_k = \begin{pmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{pmatrix}$$

and if one defines

$$H(n,m) = T(n,m)\ (J_{m+1} \otimes J_{n+1}), \tag{44}$$

the matrix H(n,m) has a Hankel block Hankel structure. It differs however with the Hankel matrices introduced by Fornasini [13] and Kao and Chen [15] to study the 2-D realization problem. This difference will be explained in Section IV, where it will be shown that the Hankel matrix H(n,m) does not play as central a role in the characterization of rational functions $g(z,\omega) = b(z,\omega)/a(z,\omega)$ as it did in the 1-D case.

The Hankel structure of H(n,m) can be exploited to relate the partial realization problem (41) with the theory of orthogonal polynomials on the hyper real line. To do so, we consider a nonnegative weight function $\mu(x,y)$ defined on $\mathbb{R}^2$, and we denote the moments of μ by

$$h_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^i \, y^i \, \mu(x,y) dxdy \quad . \tag{45}$$

Then, the matrix

$$H(n,m) = \begin{pmatrix} H_0(n) & H_1(n) & \cdots & H_m(n) \\ H_1(n) & & \iddots & \\ \vdots & \iddots & & \\ H_m(n) & & & H_{2m}(n) \end{pmatrix} \tag{46a}$$

with

$$H_j(n) = \begin{pmatrix} h_{0j} & h_{1j} & \cdots & h_{nj} \\ h_{1j} & & \iddots & \\ \vdots & \iddots & & \\ h_{nj} & & & h_{2nj} \end{pmatrix} \tag{46b}$$

can be used to characterize the inner product of polynomials of degree less than n and m in x and y. If a(x,y) and b(x,y) are two such polynomials and if we denote by a^* and b^* the vectors obtained by reversed row by row scanning of the

coefficients of $a(x,y)$ and $b(x,y)$, we have

$$\langle a,b\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(x,y)\; b(x,y)\; \mu(x,y)\; dxdy = b^{*T}\, H(n,m)a^{*}. \qquad (47)$$

Also, if we consider

$$a(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{ij}\, x^{n-i}\, y^{m-j}$$

and if a and a^* denote the vectors of coefficients of $a(x,y)$ obtained by direct and reversed row by row scanning, one has

$$a^{*} = (J_{m+1} \otimes J_{n+1})a \qquad (48)$$

where

$$a = \begin{pmatrix} a_o \\ a_j \\ a_m \end{pmatrix} \qquad \text{and} \qquad a_j = \begin{pmatrix} a_{oj} \\ a_{ij} \\ a_{nj} \end{pmatrix} .$$

Consequently, if $a_{n,m}(x,y)$ is a monic polynomial such that $a_{n,m}(x,y) \perp x^i y^j$ for $(0,0) \leq (i,j) < (n,m)$ and if one denotes

$$\langle a_{n,m}, x^n y^m\rangle = \delta_{n,m} \quad ,$$

the vector $a^*(n,m)$ of coefficients of $a_{n,m}(x,y)$ satisfies the equation

$$H(n,m)\; a^{*}(n,m) = e(n,m) \qquad (49)$$

where $e(n,m)$ is defined as in (42). By taking (44) and (48) into account, this shows that the coefficients of $a_{n,m}(x,y)$ satisfy exactly the same equation as the coefficients of $a_{n,m}(z,\omega)$ in the partial realization problem. This observation suggests that the recursions derived by Jackson [8] for the orthogonal polynomials

$a_{n,m}(x,y)$ can be adapted to the approximation problem considered here. The only difference between these two problems is that in the study of 2-D orthogonal polynomials on $\mathbb{R}^2$, the Hankel matrix $H(n,m)$ of the moments of $\mu(x,y)$ is nonnegative definite, while this is not usually the case when $H(n,m)$ is constructed from the Markov parameters of $g(z,\omega)$. However, as we shall see, this difference does not play a role in the recursions satisfied by $a_{n,m}(z,\omega)$.

Auxiliary Solutions

One of the main features of Jackson's recursions and of the recursions that we shall present below is that they require the introduction of several auxiliary solutions. At stage (n,m), we will introduce $n + m + 1$ modified Padé approximants which will be divided into $n + 1$ horizontal approximants and $m + 1$ vertical approximants. The horizontal approximants will be denoted

$$\hat{g}(z,\omega) = k^i_{n,m}(z,\omega)/h^i_{n,m}(z,\omega) \tag{50}$$

with $0 \leq i \leq n$. They are such that

$$h^i_{n,m}(z,\omega) = z^i\omega^m + \sum_{k=n-i+1}^{n} h^i_{n,m}(k,0)z^{n-k}\omega^m + \sum_{k=0}^{n}\sum_{\ell=1}^{m} h^i_{n,m}(k,\ell)z^{n-k}\omega^{m-\ell} \tag{51}$$

is an asymetric South-west (Sw) causal filter and such that $\hat{g} = k^i_{n,m}/h^i_{n,m}$ is a modified Padé approximant of g over the domain Δ described in Figure 3a. This means that $k^i_{n,m}(z,\omega)$ and $h^i_{n,m}(z,\omega)$ satisfy, respectively, the equations

$$\Pi_{-D(n,m)}h^i_{n,m}(z,\omega)\ g(z,\omega) = k^i_{n,m}(z,\omega) \tag{52a}$$

$$\Pi_{D^+(n,m)}\ h^i_{n,m}(z,\omega)\ g(z,\omega) = \delta^i_{n,m}(z)z^{-(n+1)}\omega^{-(m+1)} \tag{52b}$$

where $\delta^{i}_{n,m}(z)$ is a polynomial in z only such that

$$\deg \delta^{i}_{n,m}(z) = n-i. \tag{53}$$

Similarly, the vertical approximants will be denoted

$$\hat{g}(z,\omega) = u^{j}_{n,m}(z,\omega)/v^{j}_{n,m}(z,\omega) \tag{54}$$

where

$$v^{j}_{n,m}(z,\omega) = z^{n}\omega^{j} + \sum_{\ell=m-j+1}^{m} v^{j}_{n,m}(0,\ell)z^{n}\omega^{m-\ell}$$

$$+ \sum_{k=1}^{n}\sum_{\ell=0}^{m} v^{j}_{n,m}(k,\ell)\; z^{n-k}\,{}^{m-\ell} \tag{55}$$

is a West-south (Ws) filter with $0 \leq j \leq m$ and where $\hat{g} = u^{j}_{n,m}/v^{j}_{n,m}$ is a modified Padé approximant of g for the geometry Δ described in Figure 3b. Therefore, the polynomials $u^{j}_{n,m}(z,\omega)$ and $v^{j}_{n,m}(z,\omega)$ satisfy the equations

$$\Pi_{-D(n,m)}\; v^{j}_{n,m}(z,\omega)\; g(z,\omega) = u^{j}_{n,m}(z,\omega) \tag{56a}$$

$$\Pi_{D^{+}(n,m)}\; v^{j}_{n,m}(z,\omega)\; g(z,\omega) = \gamma^{j}_{n,m}(\omega)\; z^{-(n+1)}\omega^{-(m+1)} \tag{56b}$$

where $\gamma^{j}_{n,m}(\omega)$ is a polynomial in ω only such that

$$\deg \gamma^{j}_{n,m}(\omega) = m - j. \tag{57}$$

The total number of auxiliary approximants is only $n + m + 1$ if one observes that

$$\hat{g}(z,\omega) = k^{n}_{n,m}(z,\omega)/h^{n}_{n,m}(z,\omega) = u^{m}_{n,m}(z,\omega)/v^{m}_{n,m}(z,\omega)$$

$$= b_{n,m}(z,\omega)/a_{n,m}(z,\omega) \tag{58}$$

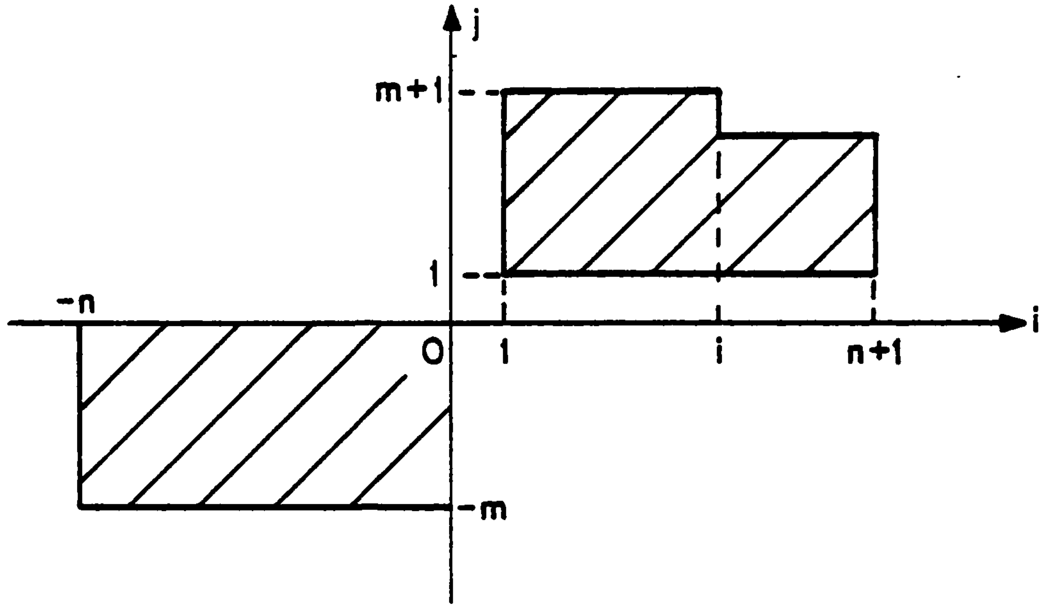

Figure 3a The Domain of Approximation of $k^i_{n,m}/h^i_{n,m}$

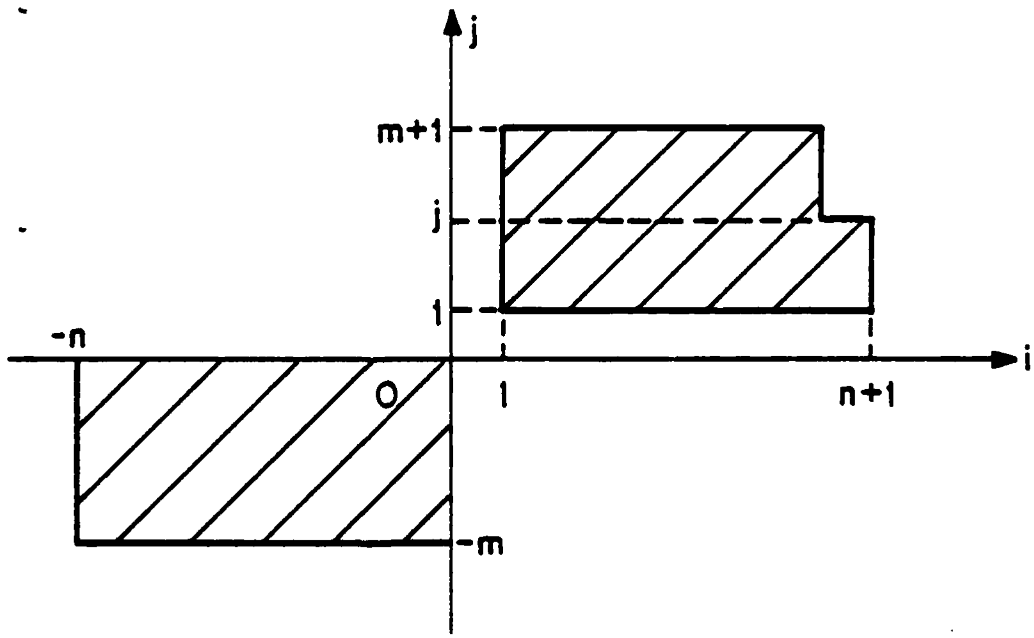

Figure 3b The Domain of Approximation of $u^j_{n,m}/v^j_{n,m}$

is the quarter plane Padé approximant that we want to compute.

The equations (52) and (56) can be written in matrix form by scanning the coefficients of $k^i_{n,m}$, $h^i_{n,m}$, $u^j_{n,m}$ and $v^j_{n,m}$ row by row and by denoting by

$$K(n,m) = (k^0(n,m), \ldots . k^n(n,m))$$

$$L(n,m) = (h^0(n,m), \ldots . h^n(n,m))$$

and

$$U(n,m) = (u^0(n,m), \ldots . u^m(n,m))$$

$$V(n,m) = (v^0(n,m), \ldots . v^m(n,m))$$

the block matrices obtained by grouping the vectors of coefficients of $k^i_{n,m}$, $h^i_{n,m}$ and $u^j_{n,m}$, $v^j_{n,m}$. Then, the numerators satisfy the relation

$$(K(n,m),\ U(n,m)) = G(n,m)\ (L(n,m),\ V(n,m)) \tag{59}$$

where G(n,m) is given by (15), so that the main problem is to compute L(n,m) and V(n,m). To do so, we shall use the identity

$$T(n,m)\ (L(n,m),\ V(n,m)) = (D(n,m),\ C(n,m)) \tag{60}$$

where T(n,m) is given by (43) and D and C are matrices of respective size $(n+1)(m+1) \times (n+1)$ and $(n+1)(m+1) \times (m+1)$ which are given by

$$D(n,m) = \left.\begin{pmatrix} 0 \\ | \\ | \\ | \\ | \\ 0 \\ 1 \end{pmatrix}\right\} m+1 \quad \otimes \quad \Delta(n,m) \tag{61a}$$

and

$$C(n,m) = \Gamma(n,m) \otimes \left.\begin{pmatrix} 0 \\ | \\ | \\ | \\ 0 \\ 1 \end{pmatrix}\right\} n+1 \quad . \tag{61b}$$

Here, $\Delta(n,m)$ and $\Gamma(n,m)$ are the $n + 1 \times n + 1$ and $m + 1 \times m + 1$ matrices whose columns $\Delta^i(n,m)$ and $\Gamma^j(n,m)$ are given, respectively, by the coefficients of the residual polynomials $\delta^i_{n,m}(z)$ and $\gamma^j_{n,m}(\omega)$. An important feature of $\Delta(n,m)$ and $\Gamma(n,m)$ is that these matrices are lower triangular. This property is a consequence of (53) and (57).

Increase in m

The introduction of the auxiliary solutions $L(n,m)$ and $V(n,m)$ is motivated mainly by the fact that these solutions can be computed recursively. To do so, we shall use an algorithm based on the block form of the recursions discussed in Lanczos [4]. The main difference with Lanczos' recursions is that when m is increased to m + 1, one has not only to compute $L(n,m+1)$, but also $V(n,m+1)$ (note that this requires the construction of one additional auxiliary solution since at stage $(n,m+1)$, there are $m + 2$ vertical approximants). Thus, we have

Computation of $L(n,m+1)$

One uses the identity

$$T(n,m+1) \left(\begin{array}{c|c|c} & 0 & 0 \\ & & 0 \\ \cline{2-3} L(n,m) & & \\ & L(n,m) & L(n,m-1) \end{array} \right) =$$

$$= \left(\begin{array}{c|c|c} \multicolumn{3}{c}{0} \\ \hline \Delta(n,m) & 0 & \Delta(n,m-1) \\ R(n,m) & \Delta(n,m) & R(n,m-1) \\ S(n,m) & R(n,m) & S(n,m-1) \end{array} \right) \tag{62}$$

which is similar to (35) with one important modification: instead of being lower triangular Toeplitz (as for $\hat{\tau}(n,m)$) the blocks of $T(n,m)$ are fully Toeplitz. This difference is significant since the algebra of Toeplitz matrices (unlike the algebra of lower triangular Toeplitz matrices) is not closed. In fact, its closure is formed by the sums of products of lower times upper Toeplitz matrices (see, e.g., [38], [39]). This means that the block entries of $L(n,m)$ as well as $\Delta(n,m)$, $R(n,m)$, $S(n,m)$ are not Toeplitz in general. These matrices are usually arbitrary. Thus, if as in (36), we define

$$M(n,m) = \Delta^{-1}(n,m-1)\ \Delta(n,m) \tag{63a}$$

and

$$N(n,m) = \Delta^{-1}(n,m)\ R(n,m) - \Delta^{-1}(n,m-1)\ R(n,m-1) \tag{63b}$$

the recursions

$$L(n,m+1) = \left(\begin{array}{c} L(n,m) \\ \hline 0 \end{array} \right) - \left(\begin{array}{c} 0 \\ \hline L(n,m) \end{array} \right) N(n,m) - \left(\begin{array}{c} 0 \\ 0 \\ \hline L(n,m-1) \end{array} \right) M(n,m) \tag{64}$$

require $0(mn)^3$ operations instead of $0(mn \log n)$ for the corresponding recursions in the geometry C of Section II.

Computation of V(n,m+1)

The first step is to observe that

$$T(n,m+1)\left(\begin{array}{c} 0 \\ \hline \\ V(n,m) \end{array}\right) L(n,m+1) = \left(\begin{array}{c|c} C(n,m) & 0 \\ & \\ \hline P(n,m) & \Delta(n,m+1) \end{array}\right) . \tag{65}$$

Therefore, if we introduce

$$Q(n,m) = \Delta^{-1}(n,m+1)\ P(n,m) \tag{66a}$$

and

$$\bar{V}(n,m+1) = \left(\begin{array}{c} 0 \\ \hline \\ V(n,m) \end{array}\right) - L(n,m+1)\ Q(n,m) , \tag{66b}$$

the matrix $\bar{V}(n,m+1)$ satisfies the equation

$$T(n,m+1)\ \bar{V}(n,m+1) = \left(\begin{array}{c} C(n,m) \\ \\ \hline 0 \end{array}\right) . \tag{67}$$

This implies that the first m + 1 columns of V(n,m+1) are given by $\bar{V}(n,m+1)$. To obtain the new auxiliary solution (i.e. the last column of V), one needs only to note from (58) that the last column of V(n,m+1) is the same as the last column of L(n,m+1). But this column has just been computed, so that

$$V(n,m+1) = (\bar{V}(n,m+1),\ h^n(n,m+1)). \tag{68}$$

The number of operations required by (68) is also $O(mn^3)$.

Increase in n

Due to the symmetry of the domain of approximation $\Delta(n,m)$, to increase n one can use the same recursions as for m, provided that the set D(n,m) is scanned column by column instead of row by row, and that the roles of L(n,m) and V(n,m) are exchanged, as well as those of n and m.

Remark 1 There is a significant difference between the previous recursions for the Padé approximants $h^i_{n,m}$ and $v^j_{n,m}$ and those derived by Jackson [8] for 2-D orthogonal polynomials on $\mathbb{R}^2$. This difference arises from the fact that there are several complete orderings of the monomials $x^i\ y^j$ which are compatible with the partial ordering

$$x^i\ y^j < x^k\ y^\ell \quad \text{iff} \quad i \le k, \quad j \le \ell.$$

The ordering considered by Jackson was based on the degree of the monomials $x^i\ y^j$, i.e.,

$$\deg x^i\ y^j = i + j$$

so that one had $x^i\ y^j < x^k\ y^\ell$ if either $\deg x^i\ y^j < \deg x^k\ y^\ell$ of if $\deg x^i\ y^j = \deg x^k\ y^\ell$ and $j < \ell$. Then, Jackson's orthogonalization procedure was based on

using the polynomials of degree n and n - 1 to compute those of degree n.

By comparison, the method that we have used to compute $h^i_{n,m}$ (resp. $v^j_{n,m}$) is based on a simple row by row (resp. column by column) lexicographic order of the monomials $z^i \omega^j$. In this framework, instead of performing a complete row by row scan of $\mathbb{N}^2$ (the rows would be infinite), we truncate the rows to length n, i.e., we consider the monomials $z^i \omega^j$ such that $i \leq n$, and we have

$$z^i \omega^j < z^k \omega^\ell$$

if either $j < \ell$ or $j = \ell$ and $i < k$. Then, the recursions for $h^i_{n,m}$ involve either an order increase ($m \to m+1$) or a change of truncation ($n \to n+1$).

Remark 2 Since $a(n,m) = h^n(n,m) = v^m(n,m)$, the previous recursions enable us to compute $a(n,m)$. To compute $b(n,m)$, one needs only to observe that $b(n,m) = k^n(n,m) = u^m(n,m)$ and that

$$(K(n,m),\ U(n,m)) = G(n,m)\ (L(n,m),\ V(n,m))$$

so that the matrices $K(n,m)$ and $U(n,m)$ obey exactly the same recursions as $L(n,m)$ and $V(n,m)$. The same observation holds for the computation of the numerator $b(n,m)$ in the geometry C of Section II.

IV. Complete Realization of 2-D Transfer Functions

The previous partial realization schemes can be related to the results obtained by Fornasini [13], Clerget [14] and Kao and Chen [15] for the study of the complete realization problem. In this case, we want to characterize the 2-D quarter-plane transfer functions $g(z,\omega)$ which are rational, i.e., such that

$$g(z,\omega) = b(z,\omega)/a(z,\omega)$$

and such that $a(z,\omega)$ has degree (n,m).

Then, if one considers the geometry B of Figure 2, for all $(p,q) \geq (n,m)$, the rational function $g(z,\omega)$ satisfies

$$\Pi_{\Delta(n,m;p,q)}\, a(z,\omega)\, g(z,\omega) - b(z,\omega) = 0 \tag{69}$$

where $\Delta(n,m;p,q)$ is defined as in (10). This means that by decomposing the domain $\Delta(n,m;p,q)$ as $\Delta(n,m;p,q) = -D(n,m) \cup \bar{\Delta}(n,m;p,q)$, where $\bar{\Delta}(n,m;p,q)$ is the complement of $-D(n,m)$ in $\Delta(n,m;p,q)$, one has

$$\Pi_{\bar{\Delta}(n,m;p,q)}\, a(z,\omega)\, g(z,\omega) = 0 \quad . \tag{70}$$

This identity can be used to characterize the transfer functions $g(z,\omega)$ which admit a rational realization of degree (n,m). To do so, we rewrite (70) in matrix form by introducing the matrices

$$T(n,m;p,q) = \begin{pmatrix} T_{m+1}(n;p) & T_m(n;p) & \cdots & T_1(n;p) \\ T_{m+2}(n;p) & T_{m+1}(n;p) & \cdots & T_2(n;p) \\ \vdots & & & \vdots \\ T_q(n;p) & T_{q-1}(n;p) & \cdots & T_{q-m}(n;p) \end{pmatrix} \tag{71a}$$

$$\tau(n,m;p) = \begin{pmatrix} T_0(n;p) & & & 0 \\ T_1(n;p) & T_0(n;p) & & \\ \vdots & \ddots & \ddots & \\ T_m(n;p) & \cdots & T_1(n;p) & T_0(n;p) \end{pmatrix} \tag{71b}$$

$$\theta(n,m;q) = \begin{pmatrix} \theta_{m+1}(n) & \theta_m(n) & \cdots & \theta_1(n) \\ \theta_{m+2}(n) & \theta_{m+1}(n) & \cdots & \theta_2(n) \\ \vdots & & & \vdots \\ \theta_q(n) & \theta_{q-1}(n) & \cdots & \theta_{q-m}(n) \end{pmatrix} \tag{71c}$$

where

$$T_j(n;p) = \begin{pmatrix} g_{n+1j} & g_{nj} & \cdots & g_{1j} \\ g_{n+2j} & g_{n+1j} & \cdots & g_{2j} \\ \vdots & & & \vdots \\ g_{pj} & g_{p-1j} & \cdots & g_{p-nj} \end{pmatrix}$$

and

$$\theta_j(n;p) = \begin{pmatrix} g_{0j} & & 0 \\ g_{1j} & g_{0j} & \\ \vdots & \ddots & \ddots \\ g_{nj} & g_{1j} & g_{0j} \end{pmatrix} .$$

Then, if we define

$$\Sigma(n,m;p,q) \triangleq \begin{pmatrix} \tau(n,m;p) \\ T(n,m;p,q) \\ \theta(n,m;q) \end{pmatrix} , \tag{72}$$

and if a denotes the vector obtained by scanning the coefficients of $a(z,\omega)$ row by row, (70) can be written as

$$\Sigma(n,m;p,q)a = 0 . \tag{73}$$

This identity leads to the following realization criterion for $g(z,\omega)$.

<u>Theorem</u> (cf. Kao and Chen [15])
If $\Sigma(n,m) \triangleq \Sigma(n,m;\ 2n+1,\ 2m+1)$, the transfer function $g(z,\omega)$ has a rational realization of order (n,m) if and only if

$$rk\Sigma(n,m) = rk\Sigma(n,m;p,q) = (n+1)(m+1) - 1 \tag{74}$$

for all $(p,q) \geq (2n+1,\ 2m+1)$, and if the first $nm + n + m$ columns of $\Sigma(n,m)$ are independent.

The main aspect of this result is that it depends on the $3(n+1)(m+1)\times(n+1)(m+1)$ matrix $\Sigma(n,m)$ which does not have either a Toeplitz or a Hankel structure (note, however, that in Kao and Chen [15], this matrix is said to be the Hankel matrix of $g(z,\omega)$). By identifying

$$\begin{aligned} \tau(n,m) &= \tau(n,m;\ 2n+1) \\ T(n,m) &= T(n,m;\ 2n+1,\ 2n+1) \end{aligned} \tag{75}$$

with the matrices defined in (24) and (43), it is clear that the matrix $T(n,m)$

(which is obtained by reordering the columns of the 2-D Hankel matrix $H(n,m)$) is only a <u>submatrix</u> of $\Sigma(n,m)$. This shows that, unlike in the 1-D case, $H(n,m)$ cannot be used to characterize the rationality of g.

V. Stochastic Modeling

In this section, the 2-D realization problem will be considered from a stochastic point of view. It will be assumed that we are given a 2-D zero-mean stochastic process $y(i,j)$, $(i,j) \in \mathbb{Z}^2$, which is space-invariant, so that its covariance is such that

$$r_{ij} = E[y(k+i, \ell+j)y(k,\ell)] \tag{76}$$

for all $(k,\ell) \in \mathbb{Z}^2$. Then, we shall seek to find a model for $y(i,j)$ which is both autoregressive and causal, so that

$$y(i,j) + \sum_{I-(0,0)} a_{k\ell} y(i-k, j-\ell) = u(i,j) \tag{77}$$

where $u(i,j)$ is a 2-D white noise, i.e.,

$$E[u(i,j)u(k,\ell)] = u\delta_{i-k}\delta_{j-\ell} \quad ,$$

and where I is a finite and causal subset of $\mathbb{Z}^2$ (by causal, we mean that I belongs to an asymmetric half plane of $\mathbb{Z}^2$). The main property of such models is that if

$$r(z,\omega) = \sum_{\mathbb{Z}^2} r_{ij} z^{-i} \omega^{-j}$$

is the spectrum of $y(\cdot,\cdot)$, and if

$$a(z,\omega) = 1 + \sum_{I-(0,0)} a_{ij} z^{-i} \omega^{-j}$$

is the polynomial associated with the filter (77), then

$$r(z,\omega) = \frac{u}{a(z,\omega)a(z^{-1},\omega^{-1})} \tag{78}$$

is a <u>spectral factorization</u> of $r(z,\omega)$. The 2-D spectral factorization problem has already been the object of a large amount of attention (cf. Helson and Lowdenslager [28], Pistor [29], Ekstrom and Woods [16], Murray [30], Delsarte, Genin and Kamp [27]), and here this problem will be considered only indirectly. However, one needs to recall the conditions of existence of factorizations such as (78). If I is chosen to be an asymmetric half-plance, i.e.,

(i) $(0,0) \in I$

(ii) $(i,j) \in I$ if and only if $(-i,-j) \notin I$, unless $(i,j) = (0,0)$

(iii) if (i,j) and $(i',j') \in I$, then $(i+j,\ i'+j') \in I$,

and if

$$\frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log r(e^{i\theta},e^{i\phi})d\theta d\phi > -\infty \quad , \tag{79}$$

it is shown in [28] that there always exists a factorization (78) such that $a(\cdot,\cdot)$ is stable. However, the support I of $a(\cdot,\cdot)$ is in general <u>infinite</u>. If $I = \mathbb{N}^2$ is the positive quarter plane, it is shown in Murray [30] that the factorization does not exist, unless the Fourier coefficients

$$s_{k\ell} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log r(e^{i\theta},\ e^{i\phi})\ \exp.\ i\ (k\theta + \ell\phi)d\theta d\phi$$

are zero in the second and fourth quadrant of $\mathbb{Z}^2$, i.e., $s_{k\ell} = 0$ if $k\ell < 0$.

In this section, we will be concerned with the problem of finding some approximants for the spectral factor $a(\cdot,\cdot)$. This approximation problem will be formulated as a causal linear least-squares estimation problem for the process $y(\cdot,\cdot)$. At stage (n,m), we shall consider the set of asymmetric half-plane

predictors of order (n,m) for $y(\cdot,\cdot)$. There are $n+m+1$ such predictors and, as in Section III, they can be divided into $n+1$ horizontal predictors $h^i_{n,m}$, $0\leq i\leq n$, and into $m+1$ vertical predictors $v^j_{n,m}$, $0\leq j\leq m$. The <u>horizontal predictors</u> are such that

$$\hat{y}(t-i,\ s|H^i_{n,m}) = -\sum_{k=i+1}^{n} h^i_{n,m}(k,0)y(t-k,\ s)$$

$$-\sum_{k=0}^{n}\ \sum_{\ell=1}^{m} h^i_{n,m}(k,\ell)y(t-k,\ s-\ell) \tag{80}$$

is the linear least-squares estimate of $y(t-i,\ s)$ given the observation of $y(\cdot,\cdot)$ over the set $H^i_{n,m}$ described in Figure 4.a (where $0\leq i\leq n$). These predictors have the property that the error

$$e(t-i,\ s|H^i_{n,m}) = y(t-i,\ s) - \hat{y}(t-i,\ s|H^i_{n,m}) \tag{81}$$

is orthogonal to $y(\cdot,\cdot)$ over the domain $H^i_{n,m}$, i.e.

$$E[e(t-i,\ s|H^i_{n,m})y(p,q)] = 0 \tag{82}$$

for all $(p,q)\ \varepsilon\ H^i_{n,m}$. By introducing the vector

$$Y(t,s;\ n,m) = \begin{pmatrix} y_0(t,s;\ n) \\ \vdots \\ y_j(t,s;\ n) \\ \vdots \\ y_m(t,s;\ n) \end{pmatrix} \tag{83a}$$

where

$$y_j(t,s;\ n) = \begin{pmatrix} y(t,\ s-j) \\ \vdots \\ y(t-i,\ s-j) \\ \vdots \\ y(t-n,\ s-j) \end{pmatrix} \tag{83b}$$

is the j^{th} row of data in $H^n_{n,m}$, and by noting that

$$e(t-i,\ s|H^i_{n,m}) = Y^T(t,s;\ n,m)h^i(n,m) \tag{84}$$

where $h^i(n,m)$ is the vector obtained by scanning the rows of the filter $h^i_{n,m}(\cdot,\cdot)$, one can rewrite (82) as

$$R(n,m)h^i(n,m) = \varepsilon^i(n,m) \tag{85}$$

where $R(n,m) = E[Y(t,s;\ n,m)Y^T(t,s;\ n,m)]$ is the covariance of $Y(t,s;\ n,m)$, and

$$\varepsilon^i(n,m) = \left.\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right\} m+1 \otimes \delta^i(n,m) \quad . \tag{86}$$

Here, $\delta^i(n,m)$ is the n+1-vector obtained by correlating $e(t-i,\ s|H^i_{n,m})$ with the row $y_0(t,s;\ n)$. This means that the last n-i entries of $\delta^i(n,m)$ are zero, but $\delta^i(n,m)$ itself is always nonzero if the prediction problem is nonsingular, i.e., if the mean-square prediction error $E[e^2(t-i,\ s|H^i_{n,m})] \neq 0$. The equation (84) presents the feature that $R(n,m)$ is Toeplitz block Toeplitz. Indeed, we have

$$R(n,m) = \begin{pmatrix} R_0(n) & R_1(n) & \cdots & R_m(n) \\ R_{-1}(n) & R_0(n) & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1(n) \\ R_{-m}(n) & \cdots & R_{-1}(n) & R_0(n) \end{pmatrix} \tag{87a}$$

with

$$R_j(n) = \begin{pmatrix} r_{0j} & r_{1j} & \cdots & r_{nj} \\ r_{-1j} & r_{0j} & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{1j} \\ r_{-nj} & \cdots & r_{-1j} & r_{0j} \end{pmatrix} \tag{87b}$$

Our main objective here will be to take advantage of this structure to obtain some recursions for the horizontal and vertical predictors $h^i_{n,m}$ and $v^j_{n,m}$.

The vertical predictors are defined symmetrically. Thus,

$$\hat{y}(t, s-j|V^j_{n,m}) = -\sum_{\ell=j+1}^{m} v^j_{n,m}(0,\ell)y(t, s-\ell)$$

$$-\sum_{k=1}^{n}\sum_{\ell=0}^{m} v^j_{n,m}(k,\ell)y(t-k, s-\ell) \tag{88}$$

is the linear least-squares estimate of $y(t, s-j)$ given the observation of $y(\cdot,\cdot)$ over the set $V^j_{n,m}$ described in Figure 4.b (where $0\leq j\leq m$). In this case,

$$e(t, s-j|V^j_{n,m}) = y(t, s-j) - \hat{y}(t, s-j|V^j_{n,m}) \tag{89}$$

is orthogonal to $y(\cdot,\cdot)$ over the domain $V^j_{n,m}$, so that by denoting by $v^j(n,m)$

the vector obtained by scanning the coefficients of $v^j_{n,m}(\cdot,\cdot)$ row by row, we obtain

$$R(n,m)v^j(n,m) = \eta^j(n,m) \quad . \tag{90}$$

The vector $\eta^j(n,m)$ is given by

$$\eta^j(n,m) = \gamma^j(n,m)\otimes \left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right)\Bigg\}n+1 \tag{91}$$

where $\gamma^j(n,m)$ is the m+1-vector obtained by correlating $e(t, s-j|V^j_{n,y})$ with the column vector

$$\bar{y}_0(t,s) = \begin{pmatrix} y(t,\ s) \\ y(t,\ s-j) \\ y(t,\ s-m) \end{pmatrix} .$$

By noting that

$$h^0(n,m) = v^0(n,m) = a(n,m) \tag{92}$$

where a(n,m) is the quarter-plane predictor of order (n,m) associated with $y(\cdot,\cdot)$, we see that the number of predictors introduced at stage (n,m) is only n+m+1. Then, if one denotes

$$H(n,m) = (h^0(n,m)\ldots h^n(n,m))$$

$$V(n,m) = (v^0(n,m)\ldots v^m(n,m)) \quad ,$$

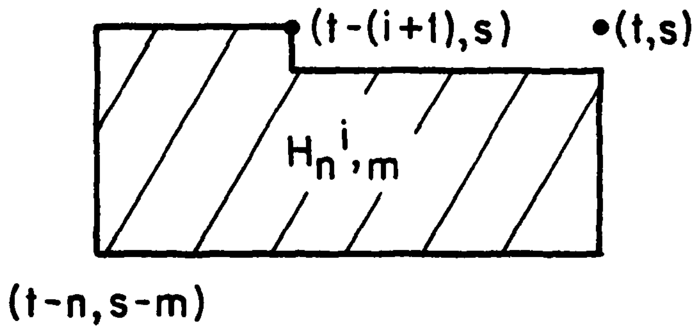

Figure 4.a The Prediction Geometry for the Horizontal Predictor $h^{i}_{n,m}$

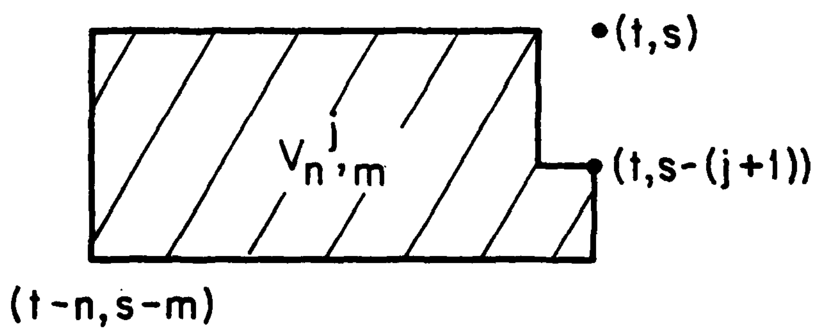

Figure 4.b The Prediction Geometry for the Vertical Predictor $v^{j}_{n,m}$

the equations (85) and (90) can be regrouped as

$$R(n,m)(H(n,m),\ V(n,m)) = (\varepsilon(n,m),\eta(n,m)) \tag{93}$$

where $\varepsilon(n,m)$ and $\eta(n,m)$ are the matrices whose columns are $\varepsilon^i(n,m)$ and $\eta^j(n,m)$. To solve (93), we shall now introduce a 2-D generalization of the recursions introduced by Levinson [19], and by Krein [40] in the continuous case, to solve 1-D Toeplitz equations.

The recursions that we consider were first presented in [20], [21] and their main feature is that n and m can be increased <u>separately</u>. By comparison, we note that the recursions introduced by Justice [22] were requiring that either n or m be fixed a priori (in fact Justice's recursions were more like those introduced by Whittle [41] and Wiggins and Robinson [42] to generalize Levinson's recursions to the 1-D matrix case). These recursions differ also from those considered by Marzetta [24] and Delsarte, Genin and Kamp [18] for asymmetric half-plane Toeplitz systems in the sense that the recursions derived by these authors were corresponding to a different geometry involving infinite vertical (or horizontal) strips.

The first step in the derivation of the 2-D Levinson recursions is to introduce the filters

$$\begin{aligned} h^{*i}_{n,m}(z,\omega) &= h^i_{n,m}(z^{-1},\omega^{-1}) \\ v^{*j}_{n,m}(z,\omega) &= v^j_{n,m}(z^{-1},\omega^{-1}) \end{aligned} \tag{94}$$

which are obtained by reversing the direction of propagation of the horizontal and vertical predictors $h^i_{n,m}$ and $v^j_{n,m}$. These filters can easily be seen to pro-

vide the linear least-squares North-east (Ne) and East-north (En) asymmetric half-plane predictors of order (n,m) associated with the process $y(\cdot,\cdot)$. Then, the vectors of coefficients of $h^{*i}_{n,m}$ and $v^{*j}_{n,m}$ are given by

$$h^{*i}(n,m) = (J_{m+1} \otimes J_{n+1})h^{i}(n,m)$$

$$v^{*j}(n,m) = (J_{n+1} \otimes J_{n+1})v^{j}(n,m) \tag{95}$$

and if one denotes by $H^{*}(n,m)$ and $V^{*}(n,m)$ the matrices obtained by regrouping these vectors, one finds that

$$R(n,m)(H^{*}(n,m),V^{*}(n,m)) = (\varepsilon^{*}(n,m), \eta^{*}(n,m)) \tag{96}$$

where

$$\varepsilon^{*}(n,m) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \otimes \Delta^{*}(n,m), \ \Delta^{*}(n,m) = J_{n+1}\Delta(n,m) \tag{97}$$

$$\eta^{*}(n,m) = \Gamma^{*}(n,m) \otimes \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \ \Gamma^{*}(n,m) = J_{m+1}\Gamma(n,m),$$

the matrices $\Delta(n,m)$ and $\Gamma(n,m)$ being the matrices whose columns are $\delta^{i}(n,m)$ and $\gamma^{j}(n,m)$. Now, the 2-D Levinson recursions can be described as follows:

<u>Increase in m</u>

To increase m, one has not only to compute H(n,m+1) (this will be done by using the block 1-D Levinson recursions), but also V(n, m+1), a step that involves the introduction of one auxiliary solution.

Computation of H(n, m+1)

Consider the identity

$$R(n,\ m+1)\begin{pmatrix} \begin{array}{c|c} H(n,m) & \begin{array}{c} 0 \\ \hline H^*(n,m) \end{array} \\ \begin{array}{c} \hline 0 \end{array} & \end{array} \end{pmatrix} = \begin{pmatrix} \begin{array}{c|c} \Delta(n,m) & \alpha^*(n,m) \\ \hline 0 & 0 \\ \hline \alpha(n,m) & \Delta^*(n,m) \end{array} \end{pmatrix} \tag{98}$$

where the residuals are

$$\alpha(n,m) \triangleq (R_{-(m+1)}(n)\ \ldots\ R_{-1}(n))H(n,m)$$

$$\alpha^*(n,m) \triangleq J_{n+1}\alpha(n,m) \quad .$$

Then, if we define

$$H(n,m+1) = \begin{pmatrix} H(n,m) \\ \hline 0 \end{pmatrix} - \begin{pmatrix} 0 \\ \hline H^*(n,m) \end{pmatrix} \rho(n,m) \tag{99}$$

with

$$\rho(n,m) = \Delta^{*-1}(n,m)\alpha(n,m) \quad , \tag{100}$$

the matrix H(n,m+1) satisfies (93), where we have

$$\Delta(n,m+1) = \Delta(n,m) - \alpha^*(n,m)\rho(n,m) \quad . \tag{101}$$

However, $\Delta(n,m+1)$ is not upper triangular in general as would be required by

the geometry of horizontal prediction chosen here. To transform it into this form, one needs only to factor $\Delta(n,m+1)$ in its upper times lower part and renormalize $H(n,m+1)$ accordingly.

Computation of V(n,m+1)

One needs to note that

$$R(n,m+1)\left(\begin{array}{c|c} 0 & \\ \hline & H(n,m+1) \\ V(n,m) & \end{array}\right) = \left(\begin{array}{c|c} \beta(n,m) & \Delta(n,m+1) \\ \hline \eta(n,m) & 0 \end{array}\right) \tag{102}$$

where

$$\beta(n,m) \triangleq (R_1(n) \; \ldots \; R_{m+1}(n))V(n,m) \quad .$$

Thus, if we introduce

$$\sigma(n,m) = \Delta^{-1}(n,m+1)\beta(n,m) \quad , \tag{103}$$

the matrix

$$\bar{V}(n,m+1) = \left(\begin{array}{c} 0 \\ \hline V(n,m) \end{array}\right) - \left(H(n,m+1)\right) \sigma(n,m) \tag{104}$$

obeys the same equation as the last m+1 columns of $V(n,m+1)$. To obtain the first column, one needs only to note that the first columns of V and H are identical, so that

$$V(n,m+1) = (h^0(n,m+1), \bar{V}(n,m+1)) \quad . \tag{105}$$

In general, as noted earlier, even though the blocks of $R(n,m)$ have a Toeplitz structure, this is not the case for $H(n,m)$ and $V(n,m)$, so that the number of operations required by the recursions (99) and (104) is $O(mn^3)$. Also, a condition for the validity of the previous recursions is that $\Delta(n,m)$ be nonsingular. But, since $\Delta(n,m)$ is upper triangular, and since its diagonal terms are equal to $E[e^2(t-i, s|H^i_{n,m})]$, $0 \leq i \leq n$, we see that $\Delta(n,m)$ will be invertible if the estimation problem is nonsingular.

Increase in n

Similar recursions can be obtained by reordering $R(n,m)$ in blocks of size $m+1 \times m+1$, and by exchanging the roles of n and m and those of H and V.

Remark. The previous recursions can be related to those obtained by Genin and Kamp [23] for 2-D orthogonal polynomials on the unit hypercircle. This relation is based on the isomorphism [43] existing between the Hilbert space of random variables $y(i,j)$, $(i,j) \in \mathbb{N}^2$ with scalar product

$$\langle y(i,j),\, y(k,\ell)\rangle = r_{i-k,j-\ell}$$

and the space of functions which are square integrable over $[0,2\pi]^2$ with respect to the positive weight function $r(e^{i\theta},e^{i\phi})$. From this point of view, the prediction problem in the plane and the one of orthogonalizing the monomials $z^i\omega^j$ with respect to $r(\cdot,\cdot)$ on the hypercircle $|z| = |\omega| = 1$ are exactly the same. However, as in the deterministic realization problem, there is a difference between our recursions and those of Genin and Kamp which is due to the total order of the monomials $z^i\omega^j$ that we have chosen. As in the deterministic case, we have chosen a lexicographic order with truncation, while Genin and Kamp's order was

based on the __grade__ of the monomials $z^i\omega^j$, i.e.,

$$\text{grade } z^i\omega^j = \max(i,j) \quad ,$$

so that one had $z^i\omega^j < z^k\omega^\ell$ if either grade $z^i\omega^j <$ grade $z^k\omega^\ell$ or grade $z^i\omega^j =$ grade $z^k\omega^\ell$ and $j < \ell$ or $k < i$.

The previous recursions enable us to compute the planar predictors $h^i_{n,m}$ and $v^j_{n,m}$ recursively. However, as was mentioned at the beginning of this section, the motivation for computing these predictors is to approximate the spectral factors of $r(z,\omega)$. We now justify this claim by considering the stability and convergence properties of $h^i_{n,m}(z,\omega)$ and $v^j_{n,m}(z,\omega)$. The first observation is that, unlike in the 1-D case, the filters $h^i_{n,m}(z,\omega)$ and $v^j_{n,m}(z,\omega)$ are __not always stable__, as noted in [23], unless $r(z,\omega)$ is separable, i.e.,

$$r(z,\omega) = r_1(z)r_2(\omega).$$

However, it was shown by Helson and Lowdenslager [28], and by Delsarte, Genin and Kamp [18], that

__Theorem: Convergence of $h^i_{n,m}$__

If the function $r(e^{i\theta},e^{i\phi})$ is strictly positive and summable over $[0,2\pi]^2$, and if $a_{Sw}(z,\omega)$ denotes the spectral factor (78) associated with the South-west Half-plane $I = \{(i,j): j < 0, \text{ or } j = 0 \text{ and } i < 0\}$, when n-k,m and $k \to \infty$, we have

$$h^k_{n,m}(e^{i\theta},e^{i\phi}) \to a_{Sw}(e^{i\theta},e^{i\phi}) \tag{106}$$

over $[0,2\pi]^2$.

Since $a_{Sw}(z,\omega)$ is stable, this shows that when n,m and i tend to infinity,

$h^i_{n,m}(z,\omega)$ is stable. In fact, it was shown by Chang and Aggarwal [17] that, for a fixed value of m, there exists an integer k_0 such that $h^i_{n,m}(z,\omega)$ is stable for $i \geq k_0$ and $n-i \geq k_0$. By symmetry, when $n,m-j \to \infty$, one gets $v^j_{n,m}(e^{i\theta},e^{i\phi}) \to a_{Ws}(e^{i\theta},e^{i\phi})$ where a_{Ws} is the spectral factor corresponding to the West-south half-plane.

However, these results do not settle completely the convergence problem for the filters $h^i_{n,m}$ or $v^j_{n,m}$. For example, the existence of limits for $h^i_{n,m}(z,\omega)$ when $n,m \to \infty$ with i constant is not clear. In the special case when $i = 0$, $h^0_{n,m}(z,\omega) = a_{n,m}(z,\omega)$ is the quarter-plane predictor of order (n,m) associated with $y(\cdot,\cdot)$ and only partial results on the convergence of $a_{n,m}(z,\omega)$ are available (see [25]-[27]). In fact, this question is related to Shanks' conjecture [44], [23] on the existence of stable planar least-squares inverses for a given 2-D polynomial, and on the existence of 2-D Beurling factorizations in the Hardy space H_2 (cf. [45]).

VI. Conclusions and Extensions

In this paper, we have presented some recursive algorithms for the deterministic and stochastic modeling of 2-D systems. These results can be extended in several directions. One such extension would be to consider the matrix form of the results presented here. We note in that respect that a matrix version of the 2-D modified approximation problem of Section III has recently been studied by Bose and Basu [46]. Also, as noted in Sections III and V, the 2-D Lanczos and Levinson recursions that we have obtained for the solution of Toeplitz block Toeplitz equations require $O(n^3m^2)$ or $O(n^2m^3)$ operations to invert a matrix with m+1 blocks of size $n+1 \times n+1$. By comparison, the 1-D Levinson algorithm requires only $O(n^2)$ operations to invert an $n+1 \times n+1$ Toeplitz matrix, so that the 2-D recursions do not preserve the efficiency of the 1-D algorithm in both dimensions. To obtain algorithms of complexity $O(n^2m^2)$ (or less), it is likely that the recent doubling algorithms of Morf [33] and Gustavson and Yun [32] will play a significant role. Another direction of generalization would be to consider 2-D extrapolation problems for the case when the process $y(\cdot,\cdot)$ is not only translation invariant, but also isotropic, i.e., when the covariance of $y(t, s)$ and $y(t', s')$ depends only on $d = ((t-t')^2 + (s-s')^2)^{1/2}$, the distance between (t,s) and (t',s'). For such processes, some extrapolation problems presenting a circular symmetry have been considered by Popov [47] and Yadrenko [48]. However, the extrapolation algorithms obtained in this context are still quite inefficient, and the problem of finding some efficient algorithms seems worth considering.

References

1. J.S.R. Chisholm, "Rational Approximants Defined from Double Power Series," Math. Comput., Vol. 17, pp. 841-848, October 1973.

2. R. Hughes Jones and G.J. Makinson, "The Generation of Chisholm Rational Approximants to Power Series in Two Variables," J. Institute Math. Applications, Vol. 13, No. 3, pp. 299-310, 1974.

3. P.R. Graves-Morris, R. Hughes Jones and G.J. Makinson, "The Calculation of Some Rational Approximants in Two Variables," J. Institute Math. Applications, Vol. 13, No. 3, pp. 310-320, 1974.

4. C. Lanczos, "Solution of Systems of Linear Equations by Minimized Iterations," J. of Res., National Bureau of Standards, Vol. 49, No. 1, pp. 33-53, July 1952.

5. E.R. Berlekamp, Algebraic Coding Theory, New York: McGraw-Hill, 1969.

6. J.L. Massey, "Shift-Register Synthesis and BCH Decoding," IEEE Trans. Inform. Theory, Vol. IT-15, No. 1, pp. 122-127, January 1969.

7. J. Rissanen, "Recursive Identification of Linear Systems," SIAM J. Control, Vol. 9, No. 3, pp. 420-430, August 1971.

8. D. Jackson, "Formal Properties of Orthogonal Polynomials in Two Variables," Duke Math. J., Vol. 2, pp. 423-434, 1936.

9. B.L. Ho and R.E. Kalman, "Effective Construction of Linear State-Variable Models from Input/Output Functions," Proc. Third Allerton Conf., pp. 449-459, 1965, and Regelungstechnik, Vol. 14, pp. 545-548, 1966.

10. L.M. Silverman, "Representation and Realization of Time-Variable Linear Systems," Department of Electrical Engineering, Columbia University, Tech. Rept. 94, June 1966.

11. D.C. Youla and P. Tissi, "N-Port Synthesis via Reactance Extraction, Part I," IEEE International Convention Record, Vol. 14, Pt. 7, pp. 183-205, 1966.

12. M. Fliess, "Matrices de Hankel," J. Math. Pures et Appliquées, Vol. 53, pp. 197-222, 1974.

13. E. Fornasini, "On the Relevance of Noncommutative Power Series in Spatial Filters Realization," IEEE Trans. Circuits Systems, Vol. CAS-25, No. 5, pp. 290-299, May 1978.

14. M. Clerget, "Realization of Two-Dimensional Recursive Filters," Proc. IFIP Congress, pp. 367-371, 1977.

15. Y.S. Kao and C.T. Chen, "Two-Dimensional Hankel Theory," to appear in Int. J. Control, 1979.

16. M.P. Ekstrom and J.W. Woods," Two-Dimensional Spectral Factorization with Applications in Recursive Digital Filtering," IEEE Trans. Acoust., Speech, Signal Processesing, Vol. ASSP-24, No. 2, pp. 115-128, April 1976.

17. Y. Chang and J.K. Aggarwal, "Design of Two-Dimensional Semicausal Recursive Filters," IEEE Trans. Circuits Syst., Vol. CAS-25, No. 12, pp. 1051-1059, December 1978.

18. Ph. Delsarte, Y. Genin and Y. Kamp, "Half-Plane Toeplitz Systems," Report R391, Philips Research Laboratory, Brussels, February 1979.

19. N. Levinson, "The Wiener RMS (Root-Mean Square) Error Criterion in Filter Design and Prediction," J. Math. Phys., Vol. 25, pp. 261-278, January 1947.

20. B.C. Lévy, S.Y. Kung and M. Morf, "New Results in 2-D Systems Theory: 2-D State-Space Models-Realization and the Notions of Controllability, Observability and Minimality," Proc. Symp. on Current Mathematical Problems in Image Science, pp. 87-95, Monterey, CA, November 1976.

21. B.C. Lévy, M. Morf and S.Y. Kung, "New Results in 2-D Systems Theory, Part III: Recursive Realization and Estimation Algorithms for 2-D Systems," Proc. 20th Midwest Symposium on Circuits and Systems, Lubbock, Texas, August 1977.

22. J.H. Justice, "A Levinson-Type Algorithm for Two-Dimensional Wiener Filtering Using Bivariate Szegö Polynomials," Proc. IEEE, Vol. 65, No. 6, pp. 882-886, June 1977.

23. Y.V. Genin and Y.G. Kamp, "Two-Dimensional Stability and Orthogonal Polynomials on the Hypercircle," Proc. IEEE, Vol. 65, No. 6, pp. 873-886, June 1977.

24. T.L. Marzetta, "A Linear Prediction Approach to Two-Dimensional Spectral Factorization and Spectral Estimation," Ph.D. Dissertation, Department of Elec. Engineering and Comp. Science, M.I.T., Cambridge, MA, January 1978.

25. Ph. Delsarte, Y.V. Genin and Y.G. Kamp, "Planar Least Squares Inverse Polynomials, Part I: Algebraic Properties," IEEE Trans. Circuits Syst., Vol. CAS-26, No. 1, pp. 59-66, January 1979.

26. Ph. Delsarte, Y. Genin and Y. Kamp, "Planar Least Squares Inverse Polynomials, Part II: Asymptotic Behavior," Report R377, Philips Research Laboratory, Brussels, July 1978.

27. Ph. Delsarte, Y. Genin and Y. Kamp, "A Survey of Two-Dimensional Toeplitz Systems," Proc. 4th Int. Symp. Mathematical Theory of Networks and Systems, Delft, The Netherlands, July 1979.

28. H. Helson and D. Lowdenslager, "Prediction Theory and Fourier Series in Several Variables, I and II," Acta Mathematica, Vol. 99, pp. 165-202, 1958 and Vol. 106, pp. 175-213, 1961.

29. P. Pistor, "Stability Criterion for Recursive Filters," IBM J. Research Devel., Vol. 18, No. 1, pp. 59-71, January 1974.

30. J.J. Murray, "Spectral Factorization and Quarter-Plane Digital Filters," IEEE Trans. Circuits Syst., Vol. CAS-25, No. 8, pp. 586-592, August 1978.

31. W.F. Trench, "An Algorithm for the Inversion of Finite Toeplitz Matrices," J. SIAM, Vol. 12, No. 3, pp. 515-522, September 1964.

32. F.G. Gustavson and D.Y.Y. Yun, "Fast Algorithms for Rational Hermite Approximation and Solution of Toeplitz Systems," IEEE Trans. Circuits Syst., Vol. CAS-26, No. 9, pp. 750-755, September 1979.

33. M. Morf, "Doubling Algorithms for Toeplitz and Related Equations," submitted to IEEE Trans. Informat. Theory.

34. R.R. Bitmead and B.D.O. Anderson, "Asymptotically Fast Solution of Toeplitz and Related Systems of Linear Equations," Tech. Report No. EE 7915, Department of Elec. Engineering, The University of Newcastle, New South Wales, Australia, July 1979.

35. S.Y. Kung, "Multivariable and Multidimensional Systems: Analysis and Design," Ph.D. Dissertation, Department of Elec. Engineering, Stanford University, Stanford, CA, June 1977.

36. T. Kailath, Introduction to Linear Systems Theory, Englewood Cliffs, N.J.: Prentice Hall, January 1980.

37. B.W. Dickinson, M. Morf and T. Kailath, "A Minimal Realization Algorithm for Matrix Sequences," IEEE Trans. Automat. Control, Vol. AC-19, No. 1, pp. 31-37, February 1974.

38. T. Kailath, S-Y. Kung and M. Morf, "Displacement Ranks of Matrices and Linear Equations," J. Math. Anal. Appl., Vol. 68, No. 2, pp. 395-407, April 1979.

39. T. Kailath, B. Lévy, L. Ljung and M. Morf, "The Factorization and Representation of Operators in the Algebra Generated by Toeplitz Operators," SIAM J. Applied Math., Vol. 37, No. 3, December 1979.

40. M.G. Krein, "The Continuous Analogues of Theorems on Polynomials Orthogonal on the Unit Circle," Dokl. Akad. Nauk SSSR, Vol. 104, pp. 637-640, 1955.

41. P. Whittle, "On the Fitting of Multivariable Autoregressions and the Approximate Canonical Factorization of a Spectral Density Matrix," Biometrika, Vol. 50, pp. 129-134, 1963.

42. R.A. Wiggins and E.A. Robinson, "Recursive Solution to the Multichannel Filtering Problem," J. Geophys. Res., Vol. 70, pp. 1885-1891, 1965.

43. T. Kailath, A. Vieira and M. Morf, "Inverses of Toeplitz Operators, Innovations and Orthogonal Polynomials, SIAM Review, Vol. 20, No. 1, pp. 106-119, 1978.

44. J.L. Shanks, S. Treitel, and J.H. Justice, "Stability and Synthesis of Two-Dimensional Recursive Filters," IEEE Trans. Audio Electroacoust., Vol. AU-20, No. 2, pp. 115-128, June 1972.

45. W. Rudin, Function Theory in Polydiscs, New York: Benjamin, 1969.

46. N.K. Bose and S. Basu, "2-D Matrix Padé Approximants: Existence, Non-Uniqueness and Recursive Computation," submitted to IEEE Trans. Automat. Control.

47. Yu. D. Popov, "On Linear Extrapolation of a Homogeneous and Isotropic Random Field by Discrete Observations on a Circle," Theory Probab. Math. Statistics, No. 4, pp. 115-120, 1974.

48. M.I. Yadrenko, "On a Linear Extrapolation Problem for an Isotropic Random Field," Theory Probab. Math. Statistics, No. 1, pp. 239-247, 1974.

Figure Captions

Figure 1. A geometry of modified approximation.

Figure 2. Some geometries of rational approximation.

Figure 3. (a) The domain of approximation of $k^{i}_{n,m}/h^{i}_{n,m}$.

(b) The domain of approximation of $u^{j}_{n,m}/v^{j}_{n,m}$.

Figure 4. (a) The prediction geometry for the horizontal predictor $h^{i}_{n,m}$.

(b) The prediction geometry for the vertical predictor $v^{j}_{n,m}$.